An interface between spatial memory and scene perception in posterior cerebral cortex

Adam Steel[1], Madeleine M. Billings[1], Edward H. Silson[2], Caroline E. Robertson[1]

[1]Department of Psychology and Brain Sciences, Dartmouth College, Hanover, NH, 03755

[2]Psychology, School of Philosophy, Psychology, and Language Sciences, University of Edinburgh, Edinburgh, UK EH8 9JZ

Corresponding author: Adam Steel, Department of Psychology and Brain Sciences, Dartmouth College, 3 Maynard Street, Hanover, NH, 03753; email: adam.steel@dartmouth.edu; tel: (202) 640 9340

Abstract word count: 75

Main text word count: 1,913

Figures: 3

Supplemental Figures: 5

Supplemental Gifs: 1

Acknowledgements: The authors would like to thank Chris Baker, Nancy Kanwisher, and Michael Cohen for helpful discussion. We also thank Tommy Botch, Yeo Bi Choi, and Anna Mynick for assistance with data collection. This work was supported by a research grant from NVIDIA to CER.

Author contribution statement: AS and CER conceived the idea and designed the research. AS collected the data. AS and MMB analyzed the data. AS wrote the initial manuscript draft. AS, EHS, and CER wrote the paper.

Conflict of interest statement: The authors declare no conflict of interest

ABSTRACT

Here, we report a new link between the spatial-memory and scene-perception systems of the human brain. Using fine-grained individual-subject fMRI, we describe three cortical areas, each lying immediately anterior to a region of the scene perception network, that are selectively activated when recalling familiar places. Network analyses show that these regions constitute a distinct functional network and interface with memory systems during naturalistic scene understanding. These regions may bridge perception and memory, supporting visually-guided navigation.

MAIN TEXT

Successful navigation requires integrating the current field-of-view with memory of the larger spatial context. While the neural systems supporting visual scene processing in posterior cerebral cortex[1–7] and spatial memory in the hippocampus and medial temporal lobe[8–15] are well-described in humans, little is known about how visual and spatio-mnemonic systems interface in the brain. Two disparate lines of inquiry yield different hypotheses. On the one hand, explicit recall of visual scenes (i.e. mental imagery) is thought to reinstate the same activity as perception in visual regions of the brain, including areas of the scene-perception network (parahippocampal place area (PPA), occipital place area (OPA), and medial place area (MPA))[16–19]. On the other hand, contemporary theories of scene processing have posited that memory-based representations are not co-localized with perceptual representations, but instead, that visual information transitions from perceptual to memory-based representations moving anteriorly from areas of the scene-perception network[1,20–23]. Distinguishing between these possibilities is critical to understanding the neural mechanisms underlying navigation. Do scene perception and place memory share common neural substrates in the human brain?

To address these issues, we used individual-subject fMRI to map the topology of scene-perception and place-memory related activity in the human brain in fine-grained detail and assess whether these responses are co-localized at the individual subject level. First, we independently localized the three regions of the scene-perception network by comparing activation when participants viewed images of scenes, people, and objects (Methods). Next, in the same subjects, we localized areas that showed preferential BOLD-activation when participants (N=14) recalled personally familiar places (e.g. their house) and people (e.g. their mother).

This analysis revealed a striking topological pattern. In all participants, we found three clusters of place-memory activity proximate to, but distinct from, the three areas of the scene-perception network, which we refer to as 'place-memory areas' for brevity (Figure 1). Across the lateral, ventral and medial cortical surfaces, the pairs of place-memory and scene-perception areas

exhibited a systematic topological relationship: the center-of-mass of each place-memory area was consistently anterior to the center-of-mass of its corresponding scene-perception area (Lateral pairs: t(13)=16.41,p<0.0001; Ventral pairs: t(13)=12.115, p<0.0001; Medial pairs: t(13)=5.855, p<0.0006; Figure 1). The anterior-shift of memory-related activation was not observed for faces: face memory was co-localized within the bounds of the perceptual FFA (t(9) = 0.1211, p = 0.906; Anterior shift for scene versus face areas, main effect ROI: F(3,85)=14.842, p<0.001, t vs faces: all ps < 0.001). We reasoned that one potential explanation for the anterior distinction of memory and perception for scene/places, but not faces might be the signal-dropout artifact affecting lateral ventral temporal cortex. However, we ruled out this explanation by replicating this experiment using an advanced multi-echo fMRI sequence that improved signal from the lateral and anterior temporal lobe and obtaining similar results (see Supplemental Material). These findings suggest that topologically separable brain areas support scene-perception and place-memory.

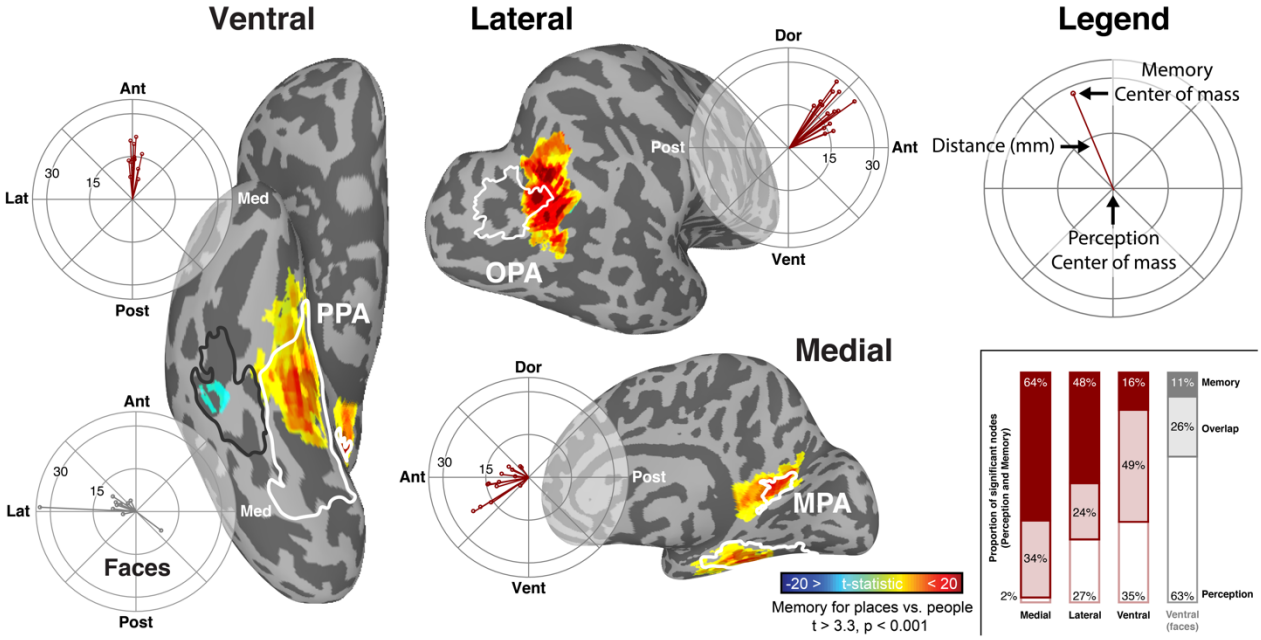


Figure 1. Distinct topology of place-memory and scene-perception responses in posterior cerebral cortex. One example participant in Experiment 1 is shown (Supplemental Figure 1 and Supplemental Gif 1 respectively show thresholded activation and unthresholded activation maps for all particpiants (n=14)). Place memory activity (warm colors) is shown along with the participant's scene perception ROIs (outlined in white). The scene-perception network (parahippocampal place area [PPA], occipital place area [OPA], and medial place area [MPA) was localized by comparing the BOLD response when participants viewed images of scenes versus with faces (outlined in white, thresholded at vertex-wise p < 0.001). Place-memory recall related areas on each surface were localized in separate fMRI runs by comparing the BOLD response when participants recalled personally familiar places versus people (warm colors, thresholded at vertex-wise p < 0.001). Scene perception and place memory activity was not co-localized. In all participants, three place-memory areas were observed, each located significantly anterior to one region of the scene-perception network. Polar plots: for each pair of regions, the center of mass of place-memory activation was significantly anterior to the center of mass of scene-perception activation in all participants (all ts > 5, p < 0.001). In contrast, face memory activation was co-localized with face perception activity on the ventral surface (cool colors, t_9 = 0.1211, p = 0.906). Inset: While the activation during place memory was systematically anterior to activation during scene perception, the spatial overlap between

perception and memory activation varied across the cortical surfaces. Note that the axes (posterior-anterior) of each polar plot is aligned to its associated cortical surface.

In addition to being topologically distinct, are the scene-perception and place-memory areas organized into distinct functional networks? In Experiment 2, we assessed the co-fluctuation of the scene-perception and place-memory areas' activity during a movie stimulus designed to elicit naturalistic scene understanding (Figure 2A). For each participant, we extracted the activity timeseries during movie watching from each ROI and calculated the correlation between the timeseries for each region pair, whilst partialling out the timeseries of all other regions. As predicted, we found that the scene-perception and place-memory areas form distinct networks: the within-network correlation was significantly greater than the between-network correlation (Main effect Network: $F(60,2)=47.99$, $p<0.0001$; Perception:Perception v Perception:Memory, $t(60)=7.69$, $p<0.0001$; Memory:Memory v Perception:Memory $t(60)=8.57$, $p<0.0001$; Memory:Memory v Perception:Perception $t(60)=1.07$, $p=0.53$; Figure 2A, B, C). We also assessed whether the scene-perception and place-memory areas preferentially associated with visual and spatio-nmemonic systems. We observed a double dissociation (Main effect of Network: $F_{(3,84)}=55.5$, $p<0.001$), whereby the scene-perception areas were more correlated to early visual cortex compared to the place-memory areas ($t_{84}=8.35$, $p<0.001$), while the place-memory areas were more strongly correlated with the hippocampus ($t_{84}=9.82$, $p<0.001$). This supports the conclusion that the scene-perception areas and place-memory areas segregate into a separable functional networks that integrate closely with the brain's visual and spatio-mnemonic systems, respectively, during scene understanding.

To confirm that the place-memory areas are preferentially driven by memory tasks, and not simply by the top-down component of the place recall task used in Experiment 1, in Experiment 3, we compared responses of the place-memory areas with the scene-perception areas when viewing panning movies of personally familiar and unfamiliar places (from Google StreetView). We found that, the place-memory areas activated significantly more than the scene-perception areas when participants saw familiar compared to unfamiliar places (Region x Familiarity interaction – Lateral: $F(1,91)=20.98$, $p<0.001$, $t(91)=4.581$; Ventral: $F(1,91)=7.00$, $p=0.01$, $t(91)=2.646$; Medial: $F(1,91)=7.321$, $p=0.008$, $t(91)=2.706$; Figure 2D, E). This suggests that activity in the place-memory network generalizes across both recall and recognition memory-related tasks.

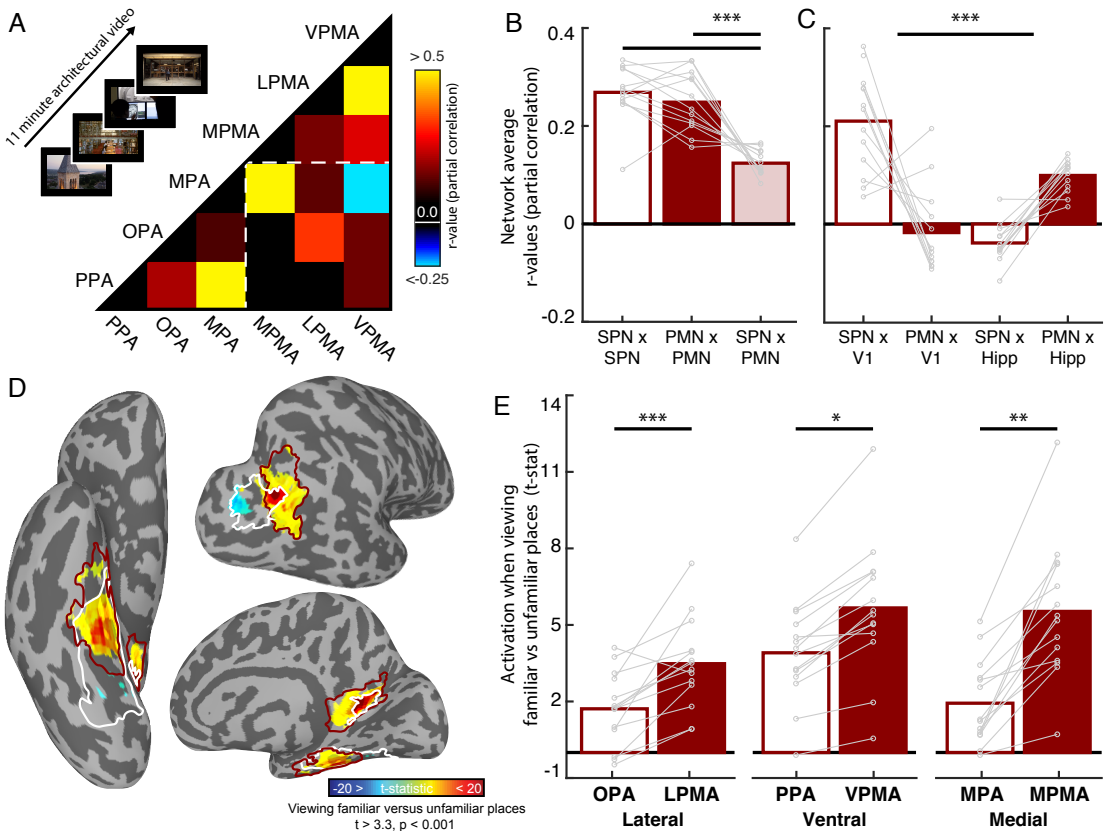


*Figure 2. The place-memory areas form a distinct functional network during scene-understanding and respond preferentially to familiar stimuli***.** *A-C. Experiment 2. To assess whether the place-memory areas and scene-perception areas form distinct functional networks, participants watched an 11-minute video designed to elicit naturalistic scene understanding, comprised of several college admissions videos, real-estate listings, and architectural tours. For each participant (n=13), the average timeseries from the scene-perception areas (parahippocampal place area [PPA], occipital place area [OPA], and medial place area [MPA], the place-memory areas (medial, ventral, and lateral place memory areas [MPMA, LPMA, VPMA]), and the hippocampus was extracted, and the pairwise partial correlation was calculated. The correlation matrix depicts the average partial correlation of each area from all participants (ordered by Ward similarity). B. The average pairwise partial correlation of within-network activity (Scene-perception network x Scene-perception network [SPN x SPN] and place-memory network x place memory network [PMN x PMN]) was significantly higher than the correlation of between network activity (SPN x PMN) ($F_{(2,60)}$=50.915, p<0.001). C. The scene-perception and place-memory areas differentially associate with the brain's visual and memory systems ($F_{(3,84)}$=55.5, p<0.001)). The scene-perception areas were more correlated with early visual cortex (t_{84}=8.35, p<0.001), while the place-memory areas were more correlated with the hippocampus (t_{84}=9.82, p<0.001), which is further evidence for their roles in perception and memory, respectively. D,E. The place-memory areas preferentially activate to familiar stimuli. D. Experiment 3. Participants viewed viewing panning movies of personally familiar places versus unfamiliar places, downloaded from Google StreetView. The cortical surface depicts the contrast of BOLD activity for a single participant, thresholded at vertex-wise p < 0.001. Only significant vertices within the scene perception (white) and place memory (burgundy) areas are shown. E. Average t-statistic of vertices in the scene-perception (open bars) and place-memory areas (filled bars) when viewing videos of personally familiar places compared to unfamiliar places. On each cortical surface, compared to the scene-perception areas, the place-memory areas showed an enhanced response to familiar stimuli (all ts > 2.6, ps < 0.01). Connected points refer to individual participants. The hippocampus also showed a preferential response to familiar compared to unfamiliar place movies (Supplemental figures 2). The amygdala (Supplemental Figure 2) and early visual cortex (Supplemental Figure 3) did not show a preferential response to familiar place movies, arguing against a purely attentional account of this effect. In all plots, error bars represent 1 SEM; n.s., p > 0.05; *p < 0.05; **p < 0.01; ***p < 0.001*

Thus far, Experiments 1-3 demonstrate that the place-memory responses in the posterior cerebral cortex are spatially separable from scene-perception areas (Exp. 1). The place-memory responsive areas co-fluctuate as a network (Exp. 2) and show enhanced recognition-related responses compared to scene-perception areas (Exp. 3). These data are strong evidence that the place-memory areas comprise a distinct functional network related to place-memory recall and recognition.

Notably, our findings conflict with a preponderance of literature suggesting explicit recall of visual scenes (mental imagery) activates the same brain areas as perception. Thus, in Experiment 4, we sought to characterize the relative roles of the scene-perception and place-memory areas in mental imagery of visual scenes, by directly comparing activation in each set of regions when participants performed mental imagery versus viewed panning movies of places (Figure 3A). As predicted from our previous results, we found that mental imagery and perception differentially engage the place-memory and scene-perception areas (Network x Task interaction – Lateral: $F(1,91)=237.37$; $p<0.001$; Ventral: $F(1,91)=78.19$; $p<0.001$; Medial: $F(1,91)=28.96$; $p<0.001$). We observed a double dissociation on all surfaces. Compared to the place-memory areas, the scene-perception areas showed greater activation during perception compared to mental imagery (Lateral: $t(91)=10.27$; $p<0.001$; Ventral: $t(91)=6.619$, $p<0.001$; Medial: $t(91)=3.289$, $p=0.008$, Figure 3B,C). In contrast, compared to the scene-perception areas, the place-memory areas were more active during mental imagery (Lateral: $t(91)=15.406$; $p<0.001$; Ventral: $t(91)=8.838$, $p<0.001$; Medial: $t(91)=5.38$, $p<0.001$; Figure 3B, C). Remarkably, PPA, a region that prior literature has suggested to be activated during mental imagery[16], showed below baseline activation during mental imagery ($t(91)=-3.132$, $p=0.008$). We also observed below-baseline activation during mental imagery in OPA ($t(13)=-2.407$, $p=0.03$). Within the scene-perception network, only MPA showed above baseline activation during both mental imagery ($t(13)=2.94$, $p=0.01$) and perception ($t(13)=4.89$, $p=0.003$). These findings directly contradict the classic understanding that mental imagery recruits the same neural substrates as perception[16–19,24], and argues that the place memory network, not the scene perception network, supports mental imagery of places.

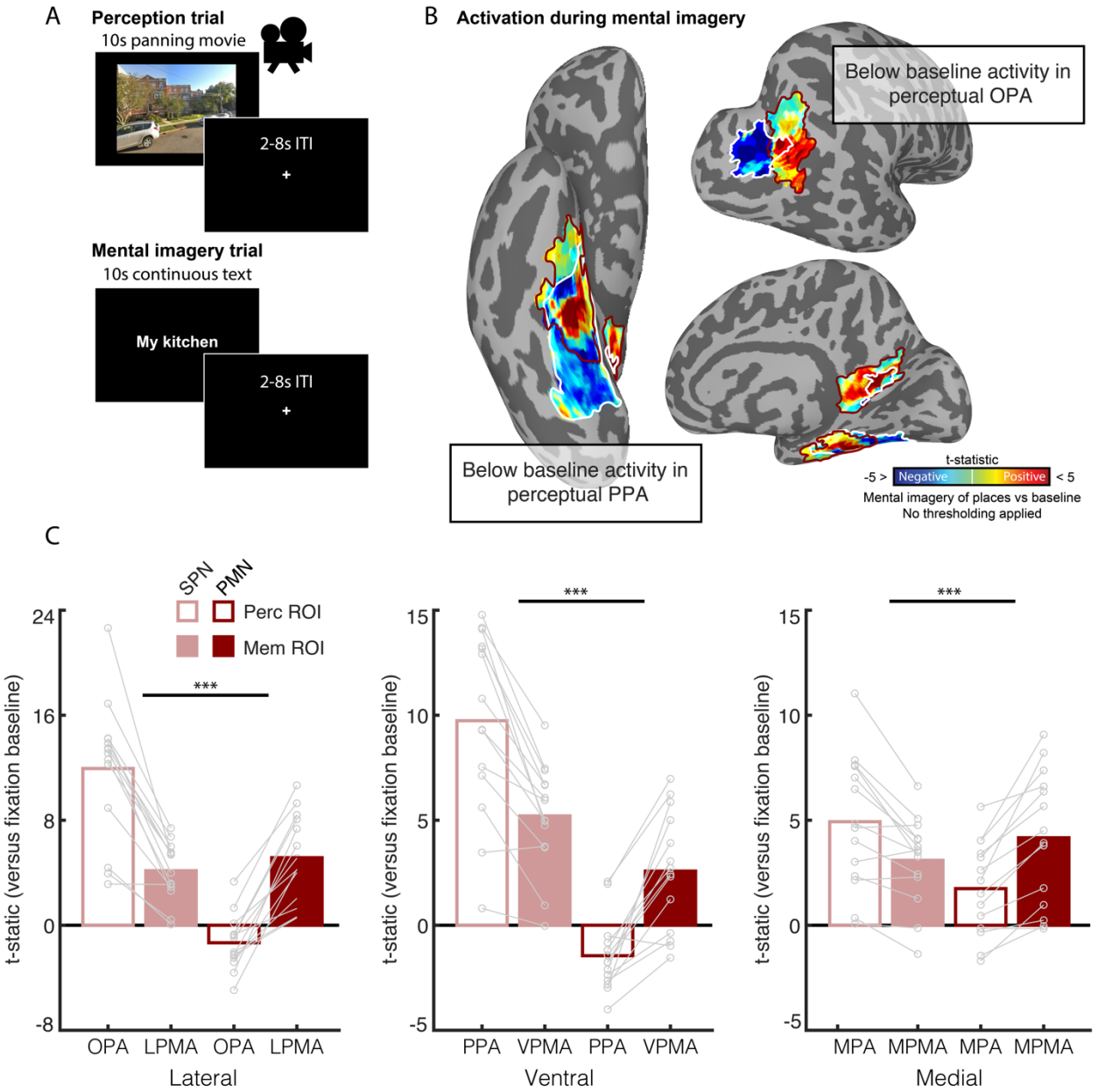


Figure 3. Perception and mental imagery differentially engage the scene-perception and place-memory areas. A. In Experiment 4, participants (n = 14) saw panning movies of unfamiliar places (perception trials) or performed mental imagery of personally familiar places (mental imagery trials). B. BOLD activation during mental imagery of places compared to baseline for a representative subject. Below baseline activation within the perceptual areas on the ventral and lateral surfaces are highlighted. Only vertices within the scene-perception network (SPN; white) and place-memory network (PMN; burgundy) are shown. C. The scene-perception areas and place-memory areas are differentially engaged during scene perception and mental imagery. Activation versus baseline of the scene-perception (open bars) and place-memory areas (filled bars) during perception of places (pink) or mental imagery of places (red). A linear mixed effects model analysis revealed that on each cortical surface, there was a significant dissociation in activation during perception and mental imagery, where the scene perception areas were significantly more active during perception, while the place memory areas were significantly more active during mental imagery of places (ROI x Task interaction – Lateral: $F_{(1,91)}$=237.37; p<0.001; Ventral: $F_{(1,91)}$=78.19; p<0.001; Medial: $F_{(1,91)}$=28.96; p<0.001). Early visual cortex (Supplemental figure 4) and amygdala (Supplemental Figure 5) also showed below baseline responses during mental

*imagery; hippocampus responded more to perception compared to imagery, but positively in both conditions (Supplemental figure 5). In all plots, error bars represent 1 SEM; n.s., $p > 0.05$; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$*

To summarize, our results suggest that scene-perception and place-memory are subserved by two distinct, but overlapping, functional networks in posterior cerebral cortex. The three regions of the place-memory network adjoin regions of the scene perception network on all three cortical surfaces – PPA (ventral), MPA (medial), and OPA (lateral) – and interface between scene perception and spatial memory systems in the human brain. Previous studies have shown that the anterior aspects of the scene perception areas, and PPA in particular, are more connected to memory structures[1,20] and harbor increasingly abstract[25] and context-related[23,26] representations relative to the posterior aspects. However, our data reveal sharp transitions from perceptual- to memory-related activity on each cortical surface, suggesting that these memory-activated regions constitute distinct brain areas. Specifically, on the ventral surface, moving posteriorly to anteriorly there is an abrupt transition from an area activated during perception (i.e. PPA) to one activated by both perception and memory. In contrast, on the medial surface, MPA is a small scene-selective visual area contained within a larger place-memory area. Only the lateral surface appears to have a transition zone between an area only activated during perception (OPA), to one activated by both perception and memory, to one activated only during memory. These transitions were evident in the unthresholded statistical maps (Supplemental Gif 1). Understanding this variable topology will likely be key to understanding these regions' functions.

Multiple aspects of our findings are surprising. First, based on previous fMRI studies, it has been widely assumed that perception and directed recall (i.e. mental imagery) of high-level stimuli (such as scenes) recruit the same neural substrates, including category-selective areas in ventral temporal cortex[16,24]. However, neuropsychiatric literature has suggested that mental imagery and visual recognition have dissociable neural substrates[27,28]. Our findings reconcile this discrepancy, by showing that distinct, but closely aligned regions subserve mental imagery and perception of places. So, why might previous fMRI studies have reported that the substrates of perception and imagery in ventral temporal cortex are shared? First, it is noteworthy that the majority of studies have used group-analysis approaches that obscure the topological relationship between perception- and memory-related activity within an individual[17,19,29]. Interestingly, early studies reported only a portion of the perceptually-activated PPA was active during imagery, but later studies using only ROI-approaches might miss this distinction[16,18,24]. Second, some studies[16] have used familiar stimuli for perception and imagery, which could recruit the anterior portion of PPA (similar to Experiment 3 here). Finally, the ability to decode stimulus category from PPA[18] during mental imagery has been taken as evidence that the same neural patterns elicited by perception are re-instantiated during memory. However, it is possible that the decodable information within PPA may arise from the nearby ventral place memory area

(VMPA). While our study was not designed for decoding analyses, we hypothesize that stimulus decoding during imagery may be localized to the place-memory areas.

Second, the place memory network described here includes portions of the caudal inferior parietal lobule, posterior parietal cortex and medial temporal lobes, which are hubs of the default mode network[30] on the lateral, medial, and ventral surfaces, respectively. These areas are known to be involved in memory processes[30–33], including autobiographical memory and semantic recall, as well as spatial memory[30–33]. While these regions may play a role in general autobiographical memory, our data show that their subregions have a specific role in spatial-memory recall and recognition (as opposed to people-memory recall, which, engages a different set of brain areas within the default mode network[21,34]). Neuropsychological patients support our assertion: damage to specific areas within the caudal inferior parietal lobule, posterior parietal cortex, or medial temporal lobe can cause severe navigational impairments, while leaving other aspects of autobiographical memory intact[35,36]. Indeed, our data suggest that memory recall is facilitated by multiple domain-specific memory areas, rather than a single general memory-recall network.

Finally, the topology of perception and memory networks was unique to the spatial domain – we found no separable networks for face-perception and people-memory recall. Why, then, might a 'place-memory network', proximate to the scene-perception network, exist, while an analogous 'face-memory network' does not? We hypothesize that the 'goal' of scene processing may be integrating the current field-of-view with the larger spatial context[37] (i.e. visual information associated with the cognitive map), which requires interfacing with memory structures like entorhinal cortex and hippocampus. In contrast, the 'goal' of visual face processing is face recognition, which requires nuanced discrimination of a specific stimulus, but not broader context[6,38,39]. Consistent with this view, portions of the scene perception network can harbor representations of spatial context beyond the current field-of-view[23,37]. All in all, we suggest that place-memory recall and mental imagery of scenes engage the place-memory network, a set of brain areas at the interface of the brain's navigational and visual systems in posterior cerebral cortex.

METHODS

Participants

Fourteen adults (9 females; age = 25.7 ± 3.6 STD years old) participated in the Experiments 1, 3, and 4. Of these, 13 participants took part in Experiment 2 (one participant requested to get out of the scanner; 8 females, age = 25.9 ±3.6 STD years old). In addition, a subset of the original participants (N=6; 2 females, age = 26.3 ±4.4 years old) completed the multi-echo replication of

Experiment 1. Participants had normal or correct-to-normal vision, were not colorblind, and were free from neurological or psychiatric conditions. Written consent was obtained from all participants in accordance with the Declaration of Helsinki and with protocol approved by the Dartmouth College Institutional Review Board (Protocol #31288). We confirmed that participants were able to perform mental imagery by assessing performance Vividness of Visual Imagery Questionaire[41]. All participants score above 9 with eyes open, indicating satisfactory performance (mean=15.2, range=9.75-19.25).

Procedure

All participants took part in Experiments 1-4. In Experiment 1, we compared the topology of activation during localizers for the category-selective scene/face perception areas with the topology of place/people memory areas collected in separate fMRI runs. In Experiment 2, we investigated the correlated activity of the place memory areas and scene perception areas during a naturalistic movie watching task to determine whether these areas constitute distinct functional networks. In Experiment 3, we compared the activation of the scene perception areas and place memory areas to panning movies of personally familiar and unfamiliar places. In Experiment 4, we compared activation in the scene perception and place memory areas to mental imagery of personally familiar places and visual perception of panning movies of unfamiliar places.

Personally familiar stimulus generation

Mental imagery stimuli

In order to perform the memory localizer (Experiment 1) and the mental imagery versus perception experiment (Experiment 4), we asked participants to generate a list of 36 familiar people and places (72 stimuli total). These stimuli were generated prior to scanning based on the following instructions:

For your scan, you will be asked to visualize people and places that are personally familiar to you. So, we need you to provide these lists for us. For personally familiar people, please choose people that you know personally (no celebrities) that you can visualize in great detail. You do not need to be in contact with these people now – just as long as you knew them personally and remember what they look like. So, you could choose your childhood friend even if you are no longer in touch with this person. Likewise, for personally familiar places, please list places that you have been and can imagine in great detail. You should choose places that are personally relevant to you, so you should avoid choosing places that you have only been one time or famous places where you have never been. You can choose places that span your whole life, so you could do your current kitchen as well as the kitchen from your childhood home.

Familiar panning movie stimuli

To generate panning movies of personally familiar places for Experiment 3, participants used the Google Maps plug-in iStreetView (istreetview.com) to locate and download 24 photospheres of personally familiar places based on the following instructions:

In another part of your scan, you will see videos of places that will be familiar or unfamiliar to you. To generate these videos, use iStreetView to locate places with which you are familiar. iStreetView has access to all photospheres that Google has access to, as well as all of the photospheres from Google Street View. So, you should do your best to find photospheres taken from perspectives where you have actually stood in your life. Like before, please find places that are personally familiar to you (such as in front of your childhood school), rather than famous places (like the Eiffel tower), even if you have been to that famous place in your life.

Participants were also asked to specify whether a particular viewpoint would be most recognizable to them, so that this field-of-view could be included in the panning movie. Once the photospheres were chosen, panning movies were generated by concatenating a series of 90° fields-of-view images of the photospheres that panning 120° to the right (in steps of 0.4°), then back to the left. The videos always began and ended at the same point of view, and the duration of the video was 10 seconds.

Category localizer - Perception

Face- and scene-selective areas were localized using a block-design functional localizer. Participants were shown images of faces, scenes, and objects (image presentation duration: 500 ms; ISI: 500 ms; block duration: 24 s; number of blocks per condition: 6). Blocks were presented continuously with no baseline periods between blocks. Participants were instructed to passively view the images and performed no task during the perceptual localizer experiment. Two runs of the localizer were performed.

Category localizer - Memory localizer

People and place selective memory areas were localized using a procedure adapted from[21]. Participants were shown written names of 36 personally familiar people and 36 places that they provided to us prior to the MRI scan. The task was performed over 4 runs, with 9 people and place stimuli presented during each run in a pseudo-randomized order. No more than 3 instances of a stimulus category (i.e. person or place) could appear consecutively.

During the task, participants were instructed to visualize the stimulus from memory as vividly as possible for the duration of the trial (10 s) following the instructions detailed below. Trials were separated by a variable inter-trial interval (4-8 s).

The instructions for mental imagery provided to the subject were as follows:

In this task, you will see the name of a person or place from the list you provided us. When the name appears, I want you to perform mental imagery of that stimulus. The mental imagery should be as vivid as possible, and you should try to do this the whole time that the name is on the screen.

If the name you see is a person, I want you to imagine that person's face in as much detail as possible. You should picture their head in front of you as the person turns their head from left to right and up and down, so that you can see all of the sides of their face. They can make different expressions, as well.

The most important thing is that you are picturing their face; when some people perform mental imagery, they will imagine interacting with that person or touching and hugging them. We do not want you to do this, just imagine the face. Also, do not picture their bodies, arms, or legs. Try to focus just on their face as it moves in your mind's eye.

If the name you see is a place, I want you to imagine that you are in the place. You should direct your attention around the place, looking to the left and right of you, and looking up and down, and behind you. We want you to have the experience of being in this place as much as possible.

When you are imagining places, you should imagine yourself standing in one single spot as you would if you were there, so don't walk around the place or imagine yourself from above. Also, we want it to be as though you are in the place, rather than thinking about the objects in the place. So, as you "look" around, don't think about the individual objects or list them in your mind. Instead, we want you to image that the place has materialized all around you, and all of the things that would be within your field of view are "visible" at the same time. For example, if you are imagining your kitchen, do not imagine the chairs, table, and windows sequentially. Instead, the experience we want you to have is as though all of these features are there at once, and that you're looking around and experiencing them as you would if you were in the room.

Between each trial, participants were instructed to clear their mind, relax, and wait for the next trial to begin.

Naturalistic movie watching experiment

To determine whether the scene perception areas and place memory areas constituted distinct functional networks, 13 participants took part in the naturalistic movie watching experiment. No audio was included in the video. We chose a natural movie watching task, rather than a simple resting-state scan, to avoid the possibility that mind-wandering or mental imagery might cause a decoupling of the perception and memory networks[31,42].

The movie consisted of an architectural tour, three college admissions videos, one tourism informational video, and one real-estate listing downloaded from YouTube. Specifically, the

segments were: 1) an architectural tour of Moosilauke Ravine Lodge in Warren, NH (0:00-2:00), 2) an admission video of Princeton University in Princeton, NJ (2:00-3:00), 3) an admission video of Dartmouth College in Hanover, NH (3:00-3:51), 4) an admission video of Cornell University in Ithaca, NY (3:52-6:40), 5) a tourism informational video of Oxford University in Oxford, United Kingdom (6:40-8:53), 6) and a real-estate listing of a lake front property in Enfield, NH (8:53-11:00). Each segment included footage of people and places, ground-level and aerial footage, and indoor and outdoor footage. Videos were chosen to match closely for content, footage type, and video quality, while varying in the degree to which participants would be familiar with the places. Participants were not asked whether places were familiar to them. The full video can be found here: URL TO VIDEO.

Familiar/unfamiliar places experiment

The goal of Experiment 3 was to confirm that the place memory areas respond preferentially to familiar stimuli (i.e. stimuli that could be recognized) presented visually. To this end, we compared brain activation when participants saw panning videos of personally familiar places with activation when viewing videos of unfamiliar places. Familiar videos were taken from photospheres as described above. Unfamiliar videos were randomly sampled from other participants. Videos of Dartmouth College and Hanover, NH were excluded from the possible unfamiliar places. All videos participant videos can be found here: URL TO VIDEO REPOSITORY.

The experiment was performed across 6 runs, with each run featuring 8 familiar and 8 unfamiliar place videos. There were 24 videos in each condition, and each video was seen twice. Before the experiment, participants were told that they would be seeing videos of familiar and unfamiliar places panning from left to right, then back again and that they should imagine that they were standing in that place, turning their head like the video. The instructions provided to the participants are reproduced below.

You will see panning movies of places that you are familiar with, as well as places that will not be familiar to you. Regardless of which you see, we want you to imagine that you are in that location and the video is your perspective as you turn your head from left to right, and back again. You should not actually turn your head.

Be sure that for all scenes you are pretending you are in that location, whether or not they are familiar to you. Of course, if it is a familiar location, you might be able to predict what is coming into view. That is fine, just do your best to imagine you're in that place for both conditions.

Each trial lasted 10 seconds. Trials were separated by a variable inter-trial interval (4-8 s). All videos were repeated two times in the experiment: once in the first three runs, and once in the last three runs.

Mental imagery versus perception experiment

Experiment 4 explicitly tested whether the place memory network was activated to a greater extent than the scene perception network during mental imagery. Each imaging run featured two types of trials, perception and mental imagery trials. On perception trials, participants saw a panning movie of an unfamiliar place, and were instructed to imagine that they were in that place turning their head (as in Experiment 3). On mental imagery trials, participants saw the name of a personally familiar place, and performed mental imagery of that place following the instructions given in Experiment 1. This task was presented over 6 runs. Each run featured 8 familiar place word stimuli for mental imagery and 8 unfamiliar place videos. As in Experiment 3, 24 unique stimuli were used in each condition (24 place word stimuli taken from the list generated in Experiment 1, and 24 panning videos of unfamiliar places). Trials lasted 10 s and were separated by a variable inter-trial interval (4-8 s). All stimuli were repeated two times in the experiment, once in the first three runs, and once in the last three runs.

MRI acquisition and preprocessing

All data was collected on a Siemens Prisma 3T MRI scanner equipped with a 32-channel head coil at Dartmouth College. Images were transformed from dicom to nifti files using dcm2niix [43], which applies slice time correction by default.

T1 image

For registration purposes, a high-resolution T1- weighted magnetization-prepared rapid acquisition gradient echo (MPRAGE) imaging sequence was acquired (TR = 2300 ms, TE = 2.32 ms, inversion time = 933 ms, Flip angle = 8°, FOV = 256 x 256 mm, slices = 255, voxel size = 1 x 1 x 1 mm). T1 images segmented and surfaces were generated using Freesurfer [44–46].

Single-echo fMRI

Acquisition

In Experiments 1-4, single-echo T2*-weighted echo-planar images covering the temporal, parietal, and frontal cortices were acquired the following parameters: TR = 2000 ms, TE = 32 ms, GRAPPA = 2, Flip angle = 75°, FOV = 240 x 240 mm, Matrix size = 80 x 80, slices = 34, voxel size = 3 x 3 x 3 mm. To minimize dropout caused by the ear canals, slices were oriented parallel to temporal lobe [47]. The initial two frames were discarded by the scanner.

Preprocessing

Task fMRI (Experiments 1, 3, and 4)

Task fMRI data was preprocessed using AFNI[48]. In addition to the frames discarded by the fMRI scanner during acquisition, the initial two frames were discarded to allow T1 signal to achieve

steady state. Signal outliers were attenuated (3dDespike). Motion correction was applied, and parameters were stored for use as nuisance regressors (3dVolreg). Data were then iteratively smoothed to achieve a uniform smoothness of 5mm FWHM (3dBlurToFWHM).

Naturalistic movie watching (Experiment 2)

Naturalistic movie watching data were preprocessed using a combination of AFNI and FSL tools. Signal outliers were attenuated (3dDespike). Motion correction was applied. Data were then iteratively smoothed to achieve a uniform smoothness of 5mm FWHM (3dBlurToFWHM).

For denoising, independent component analysis (ICA) was applied to decompose the data into signals and sources using FSL's melodic [49–51]. These were classified as signal or noise by one investigator (AS) using the approach described in Griffanti et al. (2014)[52]. Components classified as noise were then regressed out of the data (fsl_regfilt). Motion from volume registration was not included in the regression, as motion is generally well captured by the ICA decomposition. Time series were then transferred to the high-density suma standard mesh (std.141) using @SUMA_Make_Spec_FS and @Suma_AlignToExperiment.

Multi-echo fMRI

Acquisition

To better characterize activation during mental imagery of personally familiar people, we acquired a replication dataset of Experiment 1 using a multi-echo T2*-weighted sequence. Multi-echo imaging afforded the benefit of mitigating dropout with short echo times, while maintaining a high level of BOLD contrast with long echo times[53–55]. The sequence parameters were: TR = 2000 ms, TEs = [11.00, 25.33, 39.66, 53.99 ms], GRAPPA = 3, Flip angle = 75, FOV = 240 x 240 mm, Matrix size = 80 x 80, slices = 40, Multi-band factor = 2, voxel size = 3 x 3 x 3 mm. As with single-echo acquisition, the slices were oriented parallel to the temporal lobe. The initial two frames were discarded by the scanner.

Preprocessing

Multi-echo data preprocessing was implemented based on the multi-echo preprocessing pipeline from afni_proc.py. Initial preprocessing steps were carried out on each echo separately. Signal outliers were attenuated (3dDespike). Motion correction parameters were estimated from the second echo (3dVolreg); these alignment parameters were then applied to all echoes.

Data were then denoised using multi-echo ICA (meica.py[54–56]). The optimal combination of the four echoes was calculated, and the echoes were combined to form a single, optimally weighted time series (T2smap.py). Data were subsequently denoised using tedana.py. Briefly, data were subjected to PCA, and thermal noise was removed using the Kundu decision tree, which

selectively eliminates components that explain a small amount of variance and do not have a TE-dependent signal decay across echoes. Subsequently, ICA was performed to separate the time series into independent spatial components and their associated signals. These components were classified as signal and noise based on known properties of the T2* signal decay of BOLD versus noise. The retained components were then recombined to construct the optimally combined, denoised time series. Following denoising, images were blurred with a 5mm gaussian kernel (3dBlurInMask) and normalized to percent signal change.

fMRI analysis

Region of interest definitions

To define category selective perceptual areas, the scene perception localizer was modeled by fitting gamma function of the trial duration with a square wave for each condition (Scenes, Faces, and Objects) using 3dDeconvolve. Estimate motion parameters were included as additional regressors of no-interest. To compensate for slow signal drift, 4th order polynomials were included for single-echo data. Polynomial regressors were not included multiecho data analysis. Scene and face areas were drawn based on a general linear test comparing the coefficients of the GLM during scene and face trials. These contrast maps were then transferred to the suma standard mesh (std.141) using @SUMA_Make_Spec_FS and @Suma_AlignToExperiment. A vertex-wise significance of $p < 0.001$ along with expected anatomical locations was used to define the regions of interest – in particular, visually responsive PPA was not permitted to extend posterior beyond the fusiform gyrus.

To define category-selective memory areas, the familiar people/places memory data was modeled by fitting a gamma function of the trial duration for trials of each condition (people and places) using 3dDeconvolve. Estimate motion parameters were included as additional regressors of no-interest. To compensate for slow signal drift, 4th order polynomials were included for single-echo data. Polynomial regressors were not included multiecho data analysis. Activation maps were then transferred to the suma standard mesh (std.141) using @SUMA_Make_Spec_FS and @Suma_AlignToExperiment. People and place memory areas were drawn based on a general linear test comparing coefficients of the GLM for people and place memory. A vertex-wise significance threshold of $p < 0.001$ was used to draw ROIs.

For analysis of correlated activity patterns during movie watching (Experiment 2) and activation in Experiments 3 and 4, ROIs were constrained to unique members of the top 300 vertices from the perception and memory areas. These ROIs are referred to as “constrained ROIs” in subsequent sections.

Analysis of topology of perception and memory areas

The topology of the perception and memory areas was compared in two ways: by identifying whether there was a significant anterior shift from perception to memory and quantifying the overlap between category selective perception and memory vertices at the significant threshold $p < 0.001$.

To quantify the anterior displacement of perception and memory areas, we calculated the center of mass for the scene-/place-selective area on each surface, where the contribution of each vertex to the center of mass was weighted by its selectivity (t-statistic). The distance between the center of mass for perception and memory in the y-dimension (posterior-anterior) was then compared using a linear mixed effects model with ROI (perception/memory) and Hemisphere (lh/rh) as factors separately for each surface. Because there was no significant effect of hemisphere, the data are presented collapsed across hemisphere. In addition, to determine whether the anterior displacement of memory compared to perception was specific to scenes, we compared the distance in the y direction between the PPA and ventral place memory area (VPMA) with the FFA and the face-memory selectivity are on the ventral surface using a linear mixed effects model with Category (scene/face), ROI (perception/memory) and Hemisphere (lh/rh) as factors.

To further examine the relationship between perception and memory on each surface, we then qualitatively compared the overlap between the category-selective perception and memory areas on each surface.

Analysis of correlated activity timeseries during movie watching

For analysis of the co-fluctuation of activity patterns during movie watching, for each participant, we first extracted the average time course of each constrained ROI, as well as the hippocampus, which was anatomically defined based on each participant's FreeSurfer segmentation. We then calculated the correlation of the time series from each region pair while partialing out the time series from all other region pairs. The correlation matrices were calculated for each hemisphere separately. The average pairwise correlation within- and between- networks (perception: PPA, MPA, OPA; memory: VPMA, medial place memory area (MPMA), and lateral place memory area (LPMA)) were then compared using a linear mixed effects model with Connection (PxP, MxM, and PxM) and Hemisphere (lh/rh) as factors. The average correlation of each network with hippocampus was compared using a linear mixed effects model[57] with Network (Perception/Memory) and Hemisphere (lh/rh) as factors. Significant model terms were determined using an analysis of variance implemented in R[58]. Post-hoc comparisons were performed using the emmeans package[59]. There was no significant effect of hemisphere in any test, and thus data are presented collapsed across hemisphere.

Activation analysis (Experiments 3 & 4)

For Experiments 3 and 4, a gamma function of duration 10s was used to model responses for trials in each condition (Experiment 3: familiar/unfamiliar places videos; Experiment 4: mental imagery/perception). These regressors, along with motion parameters and 4th order polynomials were fit using 3dDeconvolve. Activation maps were then transferred to the SUMA standard mesh (std.141) using @SUMA_Make_Spec_FS and @SUMA_AlignToExperiment.

For analysis, the average t-statistic (compared to baseline) for each condition (Experiment 3: familiar versus unfamiliar videos; Experiment 4: mental imagery versus perception tasks) from the constrained ROIs was calculated. The average t-statistics were compared using a linear mixed effects model[57] in R[58] for each surface separately (i.e. PPA v VPMA, MPA v MPMA, OPA v LPMA were separately compared). Trial condition (Experiment 3 – Familiarity: familiar/unfamiliar; Experiment 4 – Task: perception/imagery) and Hemisphere (lh/rh) were included as factors. There was no significant effect of hemisphere in any test, and thus data are presented collapsed across hemisphere. Post-hoc tests were implemented using the emmeans package[59]. In addition, to determine whether a region was significantly active above or below baseline, the t-static from each ROI was compared versus zero. T-statistics were chosen to aid comparison across areas, which is typical in these studies[6].

REFERENCES

1. Epstein, R. A. & Baker, C. I. Scene Perception in the Human Brain. *Annu. Rev. Vis. Sci.* **5**, annurev-vision-091718-014809 (2019).

2. Epstein, R. & Kanwisher, N. A cortical representation the local visual environment. *Nature* **392**, 598–601 (1998).

3. Dilks, D. D., Julian, J. B., Paunov, A. M. & Kanwisher, N. The occipital place area is causally and selectively involved in scene perception. *J. Neurosci.* **33**, 1331–1336 (2013).

4. Marchette, S. A., Vass, L. K., Ryan, J. & Epstein, R. A. Outside looking in: Landmark generalization in the human navigational system. *J. Neurosci.* **35**, 14896–14908 (2015).

5. Marchette, S. A., Vass, L. K., Ryan, J. & Epstein, R. A. Anchoring the neural compass: Coding of local spatial reference frames in human medial parietal lobe. *Nat. Neurosci.* **17**, 1598–1606 (2014).

6. Silson, E. H., Chan, A. W. Y., Reynolds, R. C., Kravitz, D. J. & Baker, C. I. A retinotopic basis for the division of high-level scene processing between lateral and ventral human occipitotemporal cortex. *J. Neurosci.* **35**, 11921–11935 (2015).

7. Kornblith, S., Cheng, X., Ohayon, S. & Tsao, D. Y. A network for scene processing in the macaque temporal lobe. *Neuron* **79**, 766–781 (2013).

8. O'Keefe, J. & Nadel, L. *The Hippocampus as a Cognitive Map*. (Oxford University Press, 1978).

9. Moser, E. I., Kropff, E. & Moser, M.-B. Place Cells, Grid Cells, and the Brain's Spatial Representation System. *Annu. Rev. Neurosci.* **31**, 69–89 (2008).

10. Doeller, C. F., Barry, C. & Burgess, N. Evidence for grid cells in a human memory network. *Nature* **463**, 657–661 (2010).

11. Ekstrom, A. D. *et al.* Cellular networks underlying human spatial navigation. *Nature* **425**, 184–187 (2003).

12. Maguire, E. A., Woollett, K. & Spiers, H. J. London taxi drivers and bus drivers: A structural MRI and neuropsychological analysis. *Hippocampus* **16**, 1091–1101 (2006).

13. Howard, L. R. *et al.* The hippocampus and entorhinal cortex encode the path and euclidean distances to goals during navigation. *Curr. Biol.* **24**, 1331–1340 (2014).

14. Julian, J. B., Keinath, A. T., Frazzetta, G. & Epstein, R. A. Human entorhinal cortex represents visual space using a boundary-anchored grid. *Nat. Neurosci.* **21**, 191–194 (2018).

15. Tsitsiklis, M. *et al.* Single-Neuron Representations of Spatial Targets in Humans. *Curr. Biol.* **30**, 245-253.e4 (2020).

16. O'Craven, K. M. & Kanwisher, N. Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J. Cogn. Neurosci.* **12**, 1013–1023 (2000).

17. Pearson, J. The human imagination: the cognitive neuroscience of visual mental imagery. *Nat. Rev. Neurosci.* 1–11 (2019). doi:10.1038/s41583-019-0202-9

18. Boccia, M. *et al.* The dynamic contribution of the high-level visual cortex to imagery and perception. *Hum. Brain Mapp.* **40**, 2449–2463 (2019).

19. Dijkstra, N., Bosch, S. E. & van Gerven, M. A. J. Shared Neural Mechanisms of Visual Perception and Imagery. *Trends in Cognitive Sciences* **23**, 423–434 (2019).

20. Baldassano, C., Esteva, A., Fei-Fei, L. & Beck, D. M. Two Distinct Scene-Processing Networks Connecting Vision and Memory. *eNeuro* **3**, (2016).

21. Silson, E. H., Steel, A., Kidder, A., Gilmore, A. W. & Baker, C. I. Distinct subdivisions of human medial parietal cortex support recollection of people and places. *Elife* **8**, (2019).

22. Silson, E. H., Steel, A. D. & Baker, C. I. Scene-Selectivity and Retinotopy in Medial Parietal Cortex. *Front. Hum. Neurosci.* **10**, 412 (2016).

23. Peer, M., Ron, Y., Monsa, R. & Arzy, S. Processing of different spatial scales in the human brain. *Elife* **8**, (2019).

24. Ishai, A., Ungerleider, L. G. & Haxby, J. V. Distributed Neural Systems for the Generation of Visual Images. *Neuron* **28**, 979–990 (2000).

25. Baumann, O. & Mattingley, J. B. Functional organization of the parahippocampal cortex: Dissociable roles for context representations and the perception of visual scenes. *J. Neurosci.* **36**, 2536–2542 (2016).

26. Aminoff, E. M., Kveraga, K. & Bar, M. The role of the parahippocampal cortex in cognition. *Trends in Cognitive Sciences* **17**, 379–390 (2013).

27. Bartolomeo, P. The relationship between visual perception and visual mental imagery: A reappraisal of the neuropsychological evidence. *Cortex* **38**, 357–378 (2002).

28. Behrmann, M., Moscovitch, M. & Winocur, G. Intact Visual Imagery and Impaired Visual Perception in a Patient With Visual Agnosia. *J. Exp. Psychol. Hum. Percept. Perform.* **20**, 1068–1087 (1994).

29. Dijkstra, N., Bosch, S. E., Marcel, X., Van Gerven, A. J. & van Gerven, M. A. J. Vividness of Visual Imagery Depends on the Neural Overlap with Perception in Visual Areas. *J. Neurosci.* **37**, 1367–1373 (2017).

30. Andrews-Hanna, J. R., Smallwood, J. & Spreng, R. N. The default network and self-generated thought: Component processes, dynamic control, and clinical relevance. *Ann. N. Y. Acad. Sci.* **1316**, 29–52 (2014).

31. Ranganath, C. & Ritchey, M. Two cortical systems for memory-guided behaviour. *Nature Reviews Neuroscience* **13**, 713–726 (2012).

32. Renoult, L., Irish, M., Moscovitch, M. & Rugg, M. D. From Knowing to Remembering: The Semantic–Episodic Distinction. *Trends in Cognitive Sciences* **23**, 1041–1057 (2019).

33. Ritchey, M. & Cooper, R. A. Deconstructing the Posterior Medial Episodic Network. *Trends Cogn. Sci.* **xx**, (2020).

34. Peer, M., Salomon, R., Goldberg, I., Blanke, O. & Arzy, S. Brain system for mental orientation in space, time, and person. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 11072–7 (2015).

35. Takahashi, N., Kawamura, M., Shiota, J., Kasahata, N. & Hirayama, K. Pure topographic disorientation due to right retrosplenial lesion. *Neurology* **49**, 464–9 (1997).

36. Aguirre, G. K. & D'Esposito, M. Topographical disorientation: a synthesis and taxonomy. *Brain* **122**, 1613–1628 (1999).

37. Robertson, C. E., Hermann, K. L., Mynick, A., Kravitz, D. J. & Kanwisher, N. Neural Representations Integrate the Current Field of View with the Remembered 360° Panorama in Scene-Selective Cortex. *Curr. Biol.* **26**, 2463–2468 (2016).

38. Grill-Spector, K., Weiner, K. S., Kay, K. & Gomez, J. The Functional Neuroanatomy of Human Face Perception. *Annu. Rev. Vis. Sci.* **3**, 167–196 (2017).

39. Freiwald, W., Duchaine, B. & Yovel, G. Face Processing Systems: From Neurons to Real-World Social Perception. *Annu. Rev. Neurosci.* **39**, 325–346 (2016).

40. Bellmund, J. L. S., Deuker, L. & Doeller, C. F. Mapping sequence structure in the human lateral entorhinal cortex. *Elife* **8**, (2019).

41. MARKS, D. F. VISUAL IMAGERY DIFFERENCES IN THE RECALL OF PICTURES. *Br. J. Psychol.* **64**, 17–24 (1973).

42. Raichle, M. E. The Brain's Default Mode Network. *Annu. Rev. Neurosci.* **38**, 433–447 (2015).

43. Li, X., Morgan, P. S., Ashburner, J., Smith, J. & Rorden, C. The first step for neuroimaging data analysis: DICOM to NIfTI conversion. *J. Neurosci. Methods* **264**, 47–56 (2016).

44. Fischl, B. *et al.* Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron* **33**, 341–355 (2002).

45. Dale, A. M., Fischl, B. & Sereno, M. I. Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage* **9**, 179–194 (1999).

46. Fischl, B. FreeSurfer. *NeuroImage* **62**, 774–781 (2012).

47. Weiskopf, N., Hutton, C., Josephs, O. & Deichmann, R. Optimal EPI parameters for reduction of susceptibility-induced BOLD sensitivity losses: A whole-brain analysis at 3 T and 1.5 T. (2006). doi:10.1016/j.neuroimage.2006.07.029

48. Cox, R. W. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* **29**, 162–173 (1996).

49. Smith, S. M. *et al.* Advances in functional and structural MR image analysis and implementation as FSL. in *NeuroImage* **23**, S208–S219 (Academic Press, 2004).

50. Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W. & Smith, S. M. FSL. *Neuroimage* **62**, 782–790 (2012).

51. Woolrich, M. W. *et al.* Bayesian analysis of neuroimaging data in FSL. *Neuroimage* **45**, (2009).

52. Griffanti, L. *et al.* ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *Neuroimage* **95**, 232–247 (2014).

53. Poser, B. A., Versluis, M. J., Hoogduin, J. M. & Norris, D. G. BOLD contrast sensitivity enhancement and artifact reduction with multiecho EPI: Parallel-acquired inhomogeneity-desensitized fMRI. *Magn. Reson. Med.* **55**, 1227–1235 (2006).

54. Evans, J. W., Kundu, P., Horovitz, S. G. & Bandettini, P. A. Separating slow BOLD from non-BOLD baseline drifts using multi-echo fMRI. *Neuroimage* **105**, 189–197 (2015).

55. Kundu, P., Inati, S. J., Evans, J. W., Luh, W. M. & Bandettini, P. A. Differentiating BOLD and non-BOLD signals in fMRI time series using multi-echo EPI. *Neuroimage* **60**, 1759–1770 (2012).

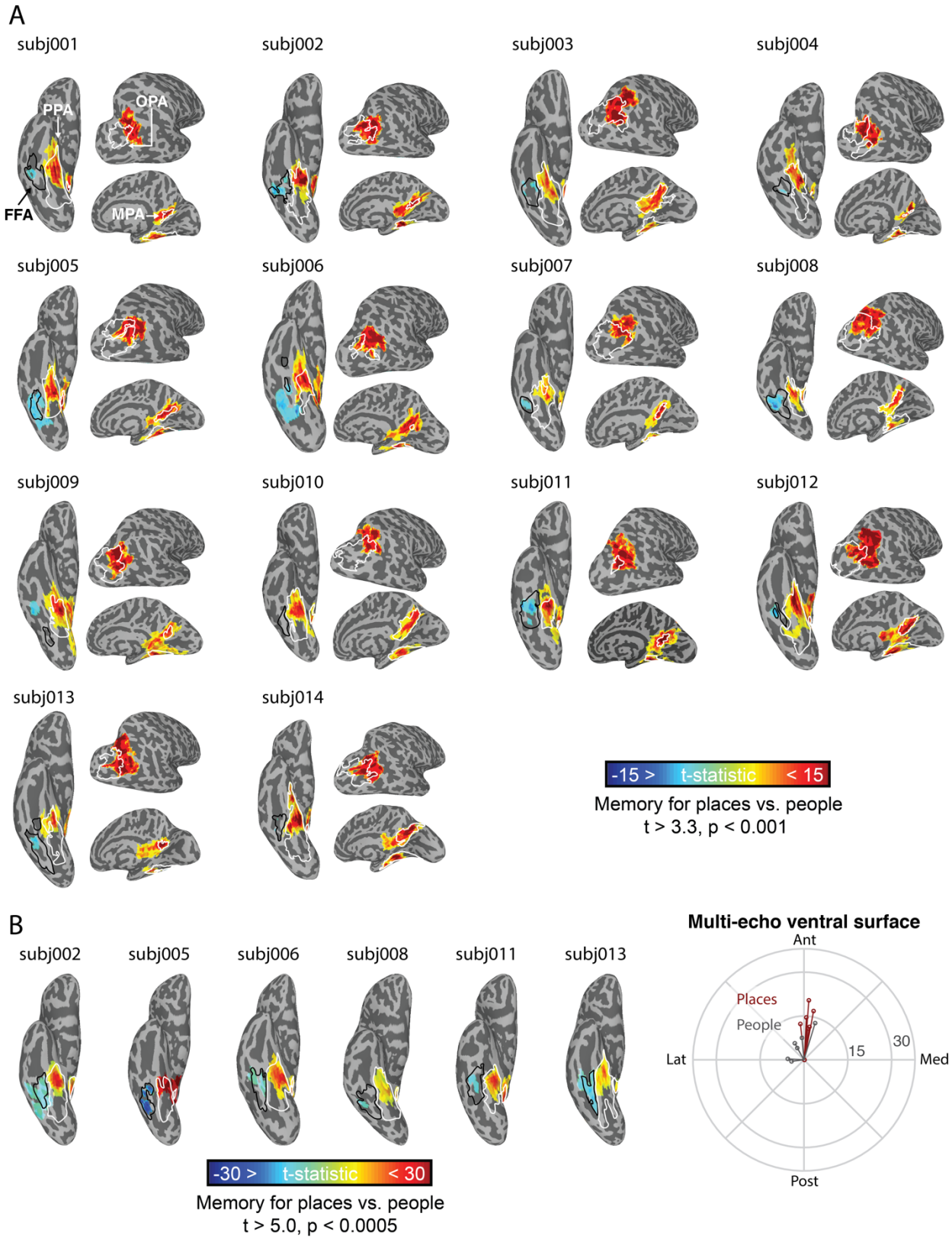
56. DuPre, E. *et al.* ME-ICA/tedana: 0.0.6. (2019). doi:10.5281/ZENODO.2558498

57. Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & Team, R. C. nlme: Linear and Nonlinear Mixed Effects Models. (2019).

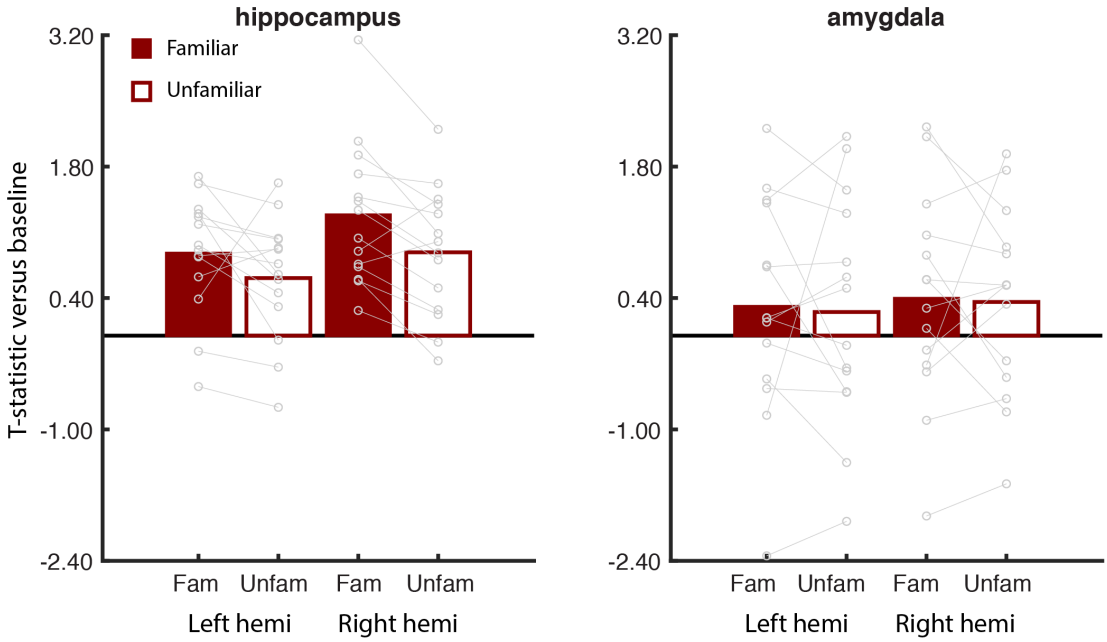
58. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computingitle. (2013).

59. Lenth, R., Singmann, H., Love, J., Buerkner, P. & Herve, M. emmeans. (2020).

Supplemental material


A
subj001
subj002
subj003
subj004
PPA
OPA
FFA
MPA
subj005
subj006
subj007
subj008
subj009
subj010
subj011
subj012
subj013
subj014
-15 > t-statistic < 15
Memory for places vs. people
t > 3.3, p < 0.001
B
subj002
subj005
subj006
subj008
subj011
subj013
-30 > t-statistic < 30
Memory for places vs. people
t > 5.0, p < 0.0005
Multi-echo ventral surface
Ant
Places
People
15
30
Lat
Med
Post

Supplemental Figure 4. Place-memory selective activation is anterior to scene-selective perceptual activation in all participants. *A. In each participant, scene-selective perceptual areas (parahippocampal place area [PPA], occipital place area [OPA], and medial place area [MPA]) were localized by comparing BOLD activation when viewing images of places compared to people); white outlined region shows each scene-selective area (PPA, OPA, MPA) thresholded at vertex-wise $p < 0.001$. Place memory selective areas were localized by comparing BOLD activation when participants recalled personally familiar places versus personally familiar people. These maps were thresholded at vertex-wise $p < 0.001$. For each participant and cortical surface, place-memory activation falls significantly anterior to scene perception activation. Face-selective perception (grey outline) and people-selective memory (cool colors) areas on the ventral surface are shown for comparison thresholded at (vertex-wise $p < 0.001$). At this threshold, people-memory activation on the ventral surface was only visible in 12/14 participants on the right hemisphere and 13/14 participants in the left hemisphere (not shown). In contrast to scene areas, the center of mass of of people-memory was not anterior to face-perception. B. Signal dropout from the air-tissue interface of the ear canal is known to obscure activity in the lateral and anterior temporal lobe, and therefore could prevent us from observing an anterior bias during people memory recall activity relative to perception of faces. To ensure this was not the case, we replicated the comparison of perception and memory localizers using multi-echo fMRI in a subset of participants, with a short-TE (11 ms) to mitigate the dropout artifact and recover signal in this area [1–3]. We were able to replicate the anterior bias in place-memory selective activation (relative to activation during scene-perception) in the replication cohort. In addition, we did not observe any consistent anterior bias for people memory activation relative to activation during face perception. All participants unthresholded activation maps for both perception and memory localizers can be found in Supplemental Gif 1.*



Supplemental Figure 5. Hippocampus responds preferentially to familiar place movies, while amygdala does not. *As an exploratory analysis, we looked to see how the subcortical memory areas responded when participants viewed familiar and unfamiliar panning movies. The average T-statistic from all voxels within the hippocampus and amygdala (defined by Freesurfer segmentation) was extracted for each participant. We then compared the activation during for each region separately using a linear mixed effects model with Hemisphere (left/right) and Familiarity (familiar/unfamiliar) as factors. The hippocampus preferentially activated to familiar compared to unfamiliar stimuli (Main effect of Familiarity: $F(1,39)=8.14$, $p = 0.0069$; $t_{39}=2.85$, $p = 0.0069$) and activation was stronger in the right hemisphere compared to the left hemisphere (Main effect of Hemisphere: $F(1,39)=8.61$, $p = 0.0056$; $t39=2.9$, $p = 0.0056$); however, there was no interaction between Hemisphere and Familiarity (Hemisphere x Familiarity interaction: $F(1,39)=0.33$, $p = 0.56$). In contrast, in the amygdala, there was no effect of Familiarity ($F(1,39)=0.29$, $p = 0.59$), Hemisphere ($F(1,39)=0.07$, $p = 0.78$), or interaction between Hemisphere and Familiarity ($F(1,39)=0.003$, $p = 0.95$). This argues against a simple account of attention modulating the response to familiar compared to unfamiliar stimuli.*

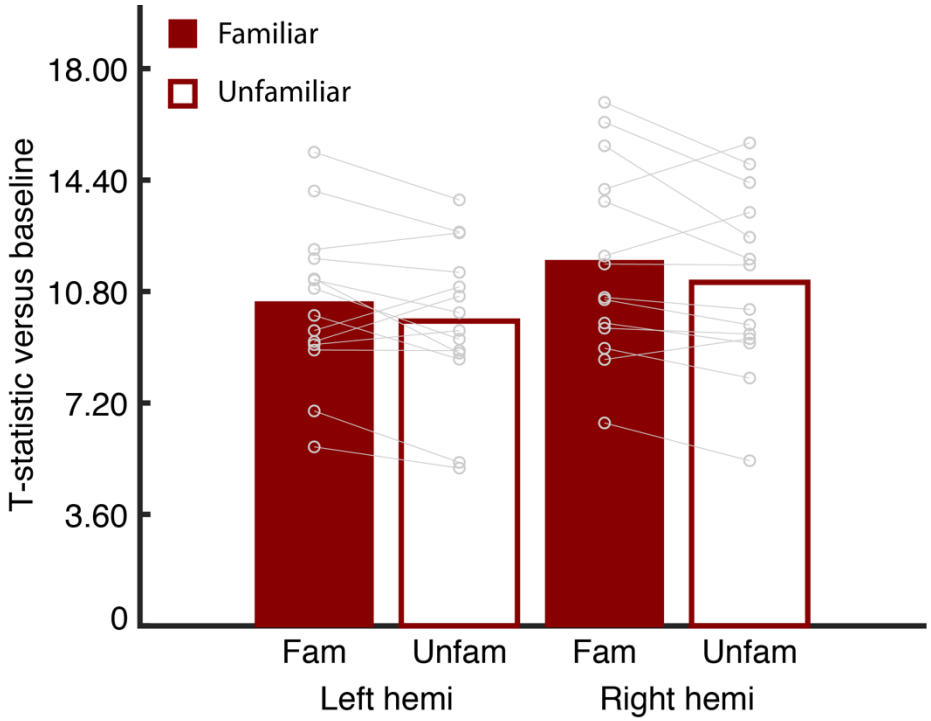


Figure 6. Early visual cortex responses to familiar and unfamiliar place movies do not differ. *One explanation for the increased activation to familiar compared to unfamiliar stimuli is that greater attention is being paid to familiar stimuli, causing a global increase in response in visually-responsive areas. To confirm that this was not the case, we compared activation of the occipital pole (defined in each participant from their Freesurfer segmentation) when participants viewed familiar versus unfamiliar panning movies. We compared the responses using a mixed effects model with Hemisphere (lh/rh) and Familiarity (familiar/unfamiliar) as factors. We found that responses in the right hemisphere were stronger than the left hemisphere (Main effect of Hemisphere: $F(1,39) = 13.29$, $p = 0.0008$; $t39=3.65$, $p = 0.0008$). However, we found no effect of Familiarity ($F(1,39)=3.19$, $p = 0.08$) or interaction between Hemisphere and Familiarity ($F(1,39)=0.01$, $p = 0.91$), confirming that a simple attentional account cannot explain the familiarity effect observed in the place memory areas.*

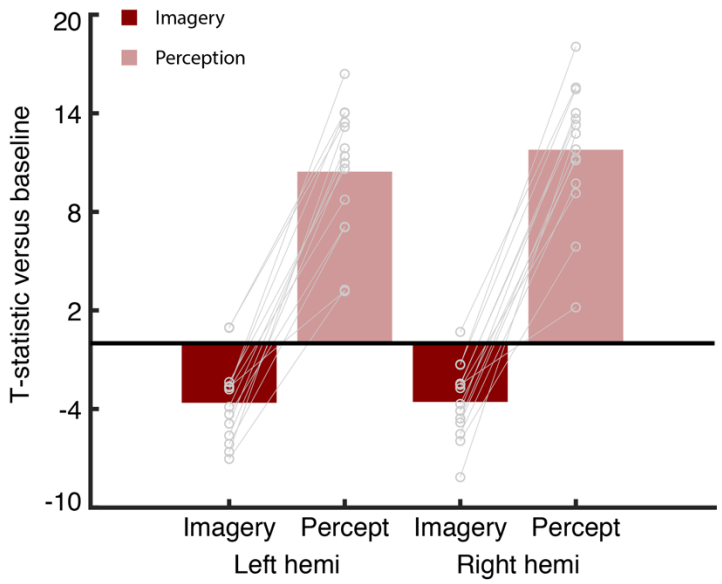


Figure 7.Early visual cortex responds more strongly to perception of place movies than mental imagery*. As an exploratory analysis, we investigated the response in early visual cortex (occipital pole defined in individual participants from their Freesurfer segmentation) during perception of unfamiliar places and mental imagery of familiar places. We compared the responses using a mixed effects model with Hemisphere (lh/rh) and Task (imagery/perception) as factors. As expected, we found that early visual cortex responded significantly more to perception compared to mental imagery (Main effect of Task: $F(1,39) = 541.38$, $p < 0.0001$; $t39 = 23.27$, $p < 0.0001$). There was no effect of Hemisphere ($F(1,39) = 1.22$, $p = 0.22$) or interaction between Hemisphere and Task ($F(1,39) = 1.05$, $p = 0.31$).*

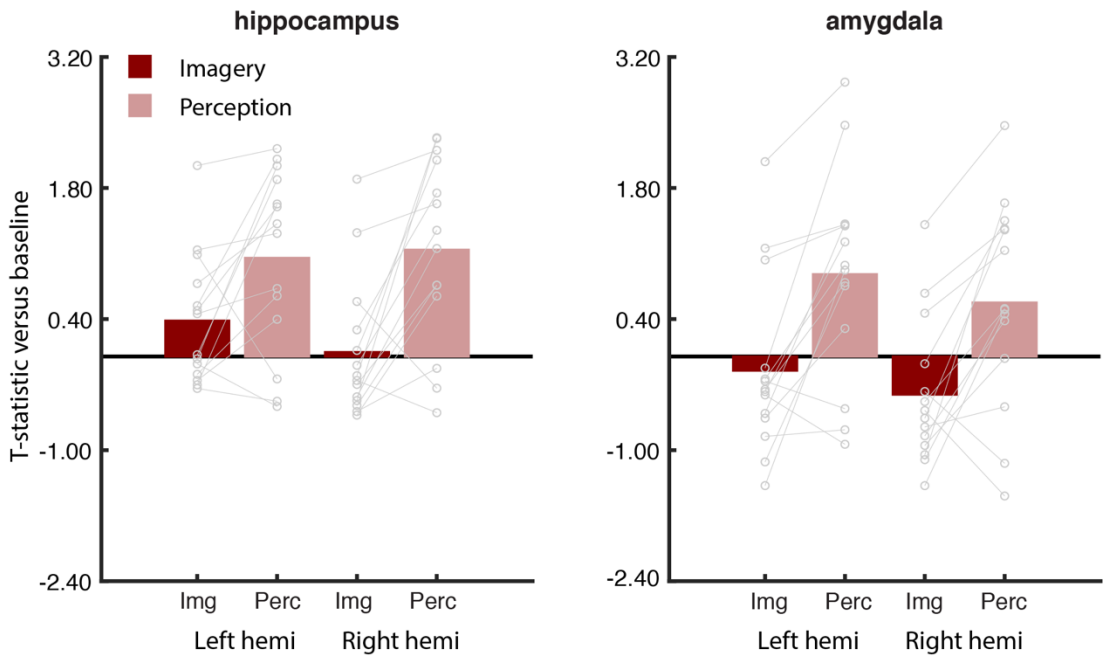


Figure 8. Both hippocampus and amygdala respond strongly during perception compared to mental imagery. *As an exploratory analysis, we investigated whether the hippocampus and amygdala — subcortical areas implicated in memory processes — responded differently when viewing panning movies of unfamiliar places and mental imagery of familiar places. The average T-statistic from all voxels within the hippocampus and amygdala (defined by Freesurfer segmentation) was extracted for each participant. We then compared the activation during for each region separately using a linear mixed effects model with Hemisphere (left/right) and Task (Imagery/Perception) as factors. Both the hippocampus and amygdala responded more strongly during perception compared to mental imagery (Hippocampus: $F(1,39)=28.67$, $p < 0.0001$; $t39=5.35$, $p < 0.0001$; Amygdala: $F(1,39)=37.22$, $p < 0.0001$; $t39=6.10$, $p < 0.0001$). Responses did not differ by hemisphere in either region (Hippocampus: $F(1,39)=0.57$, $p = 0.45$; Amygdala: $F(1,39)=2.86$, $p = 0.09$), and there was no interaction between Hemisphere and Task in either region (Hippocampus: $F(1,39)=1.65$, $p = 0.20$; Amygdala: $F(1,39)=0.02$, $p = 0.89$).*

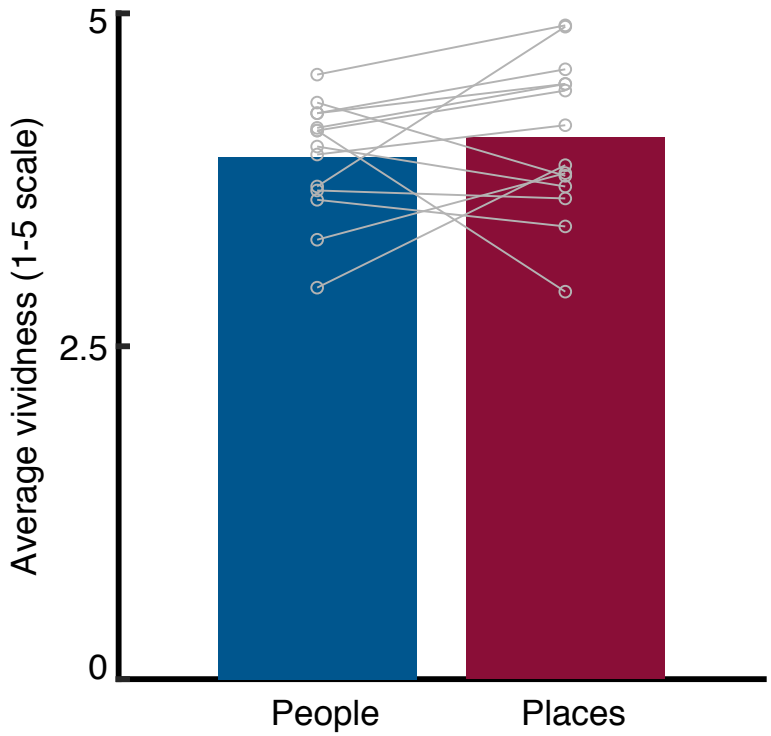


Figure 9. Participants reported no difference in vividness between imagery for people and places. *Outside of the scanner, participants rated each personally familiar stimulus for vividness of visual imagery on a scale of 1-5, where 1 indicates no imagery possible and 5 indicates "as if you were seeing the <stimulus>." There was no significant difference in vividness between people and place stimuli (t(13) = 0.92, p = 0.373), suggesting that the topological difference between place and people imagery activation extent was not due to vividness differences between the stimulus categories.*